

6 Quantitative research, agent-based modelling and theories of the social (with Yvonne Åberg)

When discussing the relationship between sociological theory and empirical research, Merton always emphasized how each draws strength from the other (e.g., Merton 1968c). Without theory, empirical research often lacks wider significance, and without empirical research, sociological theory easily turns into fictitious storytelling. Although most of us recognize the importance of a symbiotic relationship between theory and research, the current division of labour within the discipline would suggest otherwise. Most theorists specialize in theory and have little or no contact with empirical research, while empirical researchers are rarely seriously interested in theory.

In an influential article, Goldthorpe (1996) discussed how one can bridge this gap between theory and empirical research by establishing a closer link between action-based theories and quantitative research. He argued that the contribution of quantitative research to sociology 'will be seriously limited unless it is allied in some way or other to accounts of social action' (1996: 111). For a variety of reasons Goldthorpe meant that rational-choice theory was particularly well suited to this purpose. Like Edling (2000), we have a somewhat mixed attitude towards some of the details in Goldthorpe's proposal. On the one hand his arguments for establishing close links between action-based theories and quantitative research are important and to the point. On the other hand his reasons for believing that rational-choice theory is uniquely suited to integrating quantitative research and sociological theory are not as persuasive.¹ What sociology seems to need is not to bind itself to one specific substantive theory. Rather, it needs a general methodology for more closely integrating theories of the social with the results of quantitative research. As shown in this chapter, empirically calibrated agent-based models, which we refer to as 'ECA models', can accomplish this integration without imposing any *a priori* constraints on the mechanisms

¹ See chapter 3, pages 60–66 for a discussion of why rational-choice theory cannot be considered a suitable foundation for sociological theory.

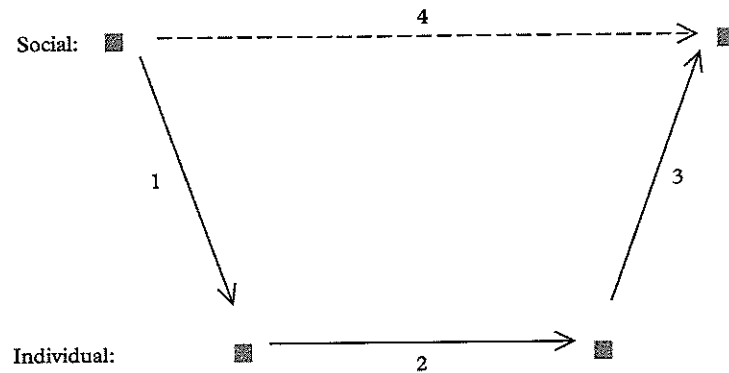


Figure 6.1. Coleman's micro-macro graph.

assumed to be operating, except, of course, that the mechanism in some way or other must be action-related. Unlike rational-choice theory, agent-based modelling is not a specific theory of action or interaction. It is a methodology for deriving the social outcomes that groups of interacting actors are likely to bring about whatever the action logics or interaction structures may be.

Coleman's (1986b) so-called micro-macro graph can be used for describing how quantitative research and agent-based modelling can complement one another (see figure 6.1). As emphasized in previous chapters, sociology is not a discipline concerned with explaining the actions or behaviours of single individuals. The focus is on larger-scale *social* phenomena characterizing groups of actors or collectivities. But the properties of these social phenomena and changes in them over time must always be explained with reference to individuals' actions, since it is individuals, not social entities, which are endowed with causal powers. Hence, even if we were exclusively interested in explaining the relationship between two social phenomena (arrow 4 in figure 6.1), a proper explanation would always entail showing how social phenomena influence individuals' actions at one point in time, and how these actions bring about the social outcomes we seek to explain at a later point in time.

As Coleman correctly pointed out, the link from the individual to the social (arrow 3) has been the main intellectual obstacle to the development of explanatory theories of the social. We know a great deal about how individuals' orientations to action, their desires, beliefs, opportunities and so forth are influenced by the social contexts in which they are embedded (arrow 1), and we also know a great deal about how their

orientations to action influence their actions (arrow 2), but when it comes to the link between individual actions and social outcomes (arrow 3) we often resort to hand-waving. This unfortunate state of affairs is due, at least in part, to the lack of an appropriate methodology for addressing these types of questions. We have a large methodological toolbox for analyzing the first two types of relations in figure 6.1, but no appropriate methodology for the third and final stage of the analysis. It is as a general methodological tool for analyzing this link between individual actions and social outcomes that agent-based modelling in general, and ECA modelling in particular, is so important for sociology.

Agent-based modelling and quantitative research have not had much influence on one another. Quantitative researchers have analyzed the first two types of relations identified in the Coleman graph without paying much attention to what these results imply for the social. When they have considered such questions they have typically ignored social interactions and assumed that the social is a simple aggregate of the individual-level entities or actions. In addition, of course, many sociologists have used time-series analyses and various forms of aggregate comparisons to try to say something about the fourth arrow in the Coleman graph. As discussed in previous chapters, however, such approaches will have little to contribute to explanatory theory because they entirely ignore the micro-level mechanisms that explain why we observe a certain change (or lack thereof) at the level of the social.

Agent-based modellers similarly have ignored much of what quantitative researchers have done and have used agent-based modelling as an exclusively theoretical tool for assessing the social outcomes that different stylized action logics and interaction structures are likely to bring about. In this chapter we seek to demonstrate how these two traditions can fruitfully complement one another. The essence of the approach advocated here is to use large-scale quantitative data to analyze and to specify the details of the first two links in the Coleman graph, and then incorporate the results of these analyses into an agent-based model in order to assess the social outcomes that are likely to be brought about (arrow 3).² We use unemployment in Stockholm during the 1990s as a

² Coleman had some ideas of his own about how one could establish a direct link between quantitative research and action-based theories, which he referred to as 'linear systems analysis' (see Coleman 1990; Coleman and Hao 1989). However, for what we believe to be good reasons, this approach never captured the attention of the sociological community. It was simply too dependent on rather implausible assumptions about the logic of action and the structure of interaction to be a useful tool for sociology in general. As far as we know, Fong (1997) is the only sociologist (in addition to Coleman and Hao) to have used the approach so far.

case study to illustrate concretely how these ideas can be put into practice.

As mentioned above, sociologists almost exclusively use agent-based modelling for theoretical purposes. Macy and Willer, for example, describe agent-based modelling as 'a new tool for *theoretical* research' (2002: 161), and they argue that the core idea behind agent-based modelling is 'to perform *highly abstract thought experiments* that explore plausible mechanisms that may underlie observed patterns' (2002: 147, emphasis added). Given the purely theoretical orientation of this field, in the next section we briefly discuss why we believe that it is important to link quantitative research and agent-based modelling to one another. Thereafter we give a substantive background to our case study, which focuses on the role of social interactions in explaining spatial and temporal variations in youth unemployment. We then use a large-scale data set to empirically specify the first two links in the Coleman graph. We use this data to estimate how individuals' likelihood of leaving unemployment are influenced by various individual-level and social-level phenomena, including the unemployment level among those with whom the individuals interact. First we use some of the estimates from these analyses to inject some realism into the type of agent-based model analyzed in chapter 4. In order to predict how the probability of leaving unemployment affects the unemployment level, we thereafter develop the ECA model, and we use this model as a virtual laboratory to examine how various changes in the micro-level processes are likely to influence the level and spatial variation in unemployment. Social processes in which large numbers of heterogeneous actors influence one another through time are rather complex. As a result of this, some of the analyses reported below are also rather complex. This is an unfortunate but unavoidable consequence of the complexity of the subject matter.

Quantitative research and agent-based modelling

Computational modelling has changed a great deal over recent decades. Macy and Willer (2002) aptly describe the general trend as representing a change from factor-based to actor-based models (see also Gilbert and Troitzsch 1999). While social simulations used to be variable-based and sought to reproduce the aggregate dynamics of social systems, the trend has been towards actor-based models. These actor-based models first took the form of so-called micro-simulations, but during the last decade agent-based models have come to dominate (e.g., Carley 1991; Epstein and Axtell 1996; Macy 1991; Mark 1998). The distinguishing feature

of an agent-based model is that it explains social phenomena from the bottom up, that is, social phenomena are analyzed as the outcomes of the actions of interacting actors.

While we consider this development towards actor-based models to be of fundamental importance for sociology as an explanatory science, it is important to recognize that this transition in most cases has also meant a change from empirically calibrated models to non-empirical models constructed by researchers to capture the logic of a particular theoretical mechanism. If we sought to derive the social-level consequences of a stylized theoretical mechanism, as was done in chapter 4 and as most agent-based modellers seek to do, this is exactly the type of model we should use. As noted above, however, agent-based modelling also is valuable for other reasons. Most importantly, it can be used for linking empirical research findings to their implied social-level consequences. When agent-based modelling is used for this latter purpose it is essential that the specification of the agent-based model is closely informed by the results of statistical analyses.

The results of such statistical analyses should influence both the ways in which the operative mechanisms are modelled and the set of confounding factors taken into account in the analysis. For example, instead of making up a rule for how actors' opinions or actions are influenced by the opinions or actions of others, as was done in chapter 4, one should use statistical analyses to arrive at a specification that as closely as possible mirrors how the relevant actors actually interacted and influenced one another. Similarly, and as discussed in chapter 4, societies are not closed systems. We must always allow for the possibility that various events or processes, external and unrelated to the processes we focus upon, may influence the outcomes we seek to explain. Not taking into account selection and environmental effects, for instance, may easily lead us astray.³ Unless we are able to distinguish between these different types of processes, in the statistical as well as in the agent-based analyses, the usefulness of the approach advocated here is much reduced.

Establishing closer links between quantitative research and agent-based modelling thus promises to accomplish two different tasks. First, it provides a test of the agent-based model in the sense that it examines the extent to which it can bring about the social outcome it seeks to explain also for realistic parameter values. Second, it provides a

³ See chapter 3, pages 45–47 for the distinctions between interaction, selection and environmental effects.

micro-macro link that allows us to derive the social-level implications of a set of quantitative research results.

We want to emphasize that we are not advocating a return to older system-level or micro-simulation techniques. The type of empirically calibrated agent-based model that we have in mind is a true agent-based model in the sense that it is a bottom-up model in which agents in interaction with one another bring about various social outcomes. But the model is calibrated with real data, and it takes into account various real-world events taking place during the course of the analysis. After giving a background to our case study and empirically assessing how important social interactions are for unemployment durations, we give precise content to these ideas by developing and analyzing an agent-based model that fulfils these requirements.

Social interactions and youth unemployment

During the 1990s unemployment figures rose sharply throughout the western world, particularly among young people. In Sweden, the focus of this empirical study, unemployment levels among young people had not been so high since the economic recessions of the 1930s. Our purpose here is not to try to explain why unemployment levels changed as they did. Instead, we focus on one specific type of mechanism that has not received sufficient attention in the literature but which nevertheless is likely to have been of considerable importance. We focus on social interactions and their potential importance in explaining temporal and spatial variations in unemployment.

Social interactions can influence unemployed individuals' actions for a variety of reasons, and in order to understand better why we observe what we observe it is essential to try to distinguish between them. As suggested in chapter 3, one should at least try to distinguish between three broad types of social interactions: opportunity-based, belief-based and desire-based. Consider the case where the focal actor is an unemployed individual and the action focused upon is one that increases the likelihood of the individual leaving the unemployed state. How can this action be influenced by the unemployment level among the individual's peers? The general answer is that this can occur in three distinct ways: (1) the unemployment level among peers can influence the focal individual's opportunities and thereby his or her choice of action; (2) it can influence the focal individual's beliefs and thereby his or her choice of action; and (3) it can influence the focal individual's desires and thereby his or her choice of action.

As observed by Granovetter (1974) and others, many individuals obtain their jobs via informal social contacts with friends and acquaintances, who pass on information about jobs to prospective job candidates and information about potential job candidates to employers. If the unemployment rate is high among friends and acquaintances, the quality of this information network is lowered and information about vacant jobs will not reach the focal actor to the same extent as if friends and acquaintances were employed. Therefore, the focal actor's probability of finding a job will be negatively influenced by the unemployment level among friends and acquaintances. This is an example of an opportunity-based interaction effect.

The individual's likelihood of leaving the unemployed state is also likely to be influenced by his or her beliefs about the jobs that he or she can expect to get. Traditional decision and search theory would suggest that those who expect to get a job, particularly a high-paying one, would invest more time and energy in a job search than those with bleaker prospects. To the extent that these beliefs are partly influenced by the experiences of friends, acquaintances or neighbours, we have an example of a belief-based interaction effect. One example of belief-based interaction is the so-called *discouraged worker effect*, that is, that a high unemployment rate may discourage individuals from looking for work because they do not expect to find any (e.g., Schweitzer and Smith 1974). Another type of belief-based interaction occurs when other individuals serve as role models for the focal individual. One reason for Wilson's concern about the exodus of middle-class families from many ghetto neighbourhoods in the United States, for example, was the influence of precisely such belief-based interaction effects: 'the very presence of these families . . . provides mainstream role models that help keep alive the perception that education is meaningful, that steady employment is a viable alternative to welfare, and that family stability is the norm, not the exception' (Wilson 1987: 56). In both the discouraged-worker and the role-model cases, unemployment among others influences the focal individual's beliefs such that his or her chances of leaving the unemployed state are altered.

There are also reasons to believe that desire-based interactions are important in this context. One such reason is the existence of the social norm, which holds that one should earn one's income. Being unemployed usually means that one cannot live up to this norm, and this may bring about feelings of shame or embarrassment (Elster 1989a). Such feelings in large part can be attributed to deviations from what is normal or typical in the unemployed individual's reference groups (Sherif and Sherif 1964). Since reference groups vary among individuals, however, the normative pressure is not likely to be felt with equal

intensity
employed
and the le
such emo
likely to b
of one's
several dif
ment amo
social and

Clark (2
ment of o
ence. Bas
that it was
ured by a
many peo
unemploy
the most c
jobs and v
Oswald 19

Clark's
of unempl
consequen
that increa
ment. The
the unemp
they accep
in searchin
likely that
have at lea
purely eco
non-pecun
the pecun
consequen

To the
employed
the unemp
leaving the
based inte

⁴ They are
automatic
self-reinfor
cesses that

4) and others, many individuals contacts with friends and acquaintances to prospective job candidates and employers. If the contacts and acquaintances, the quality and information about vacant jobs is the same extent as if friends and acquaintances, the focal actor's probability of finding a job is increased by the unemployment level. This is an example of an opportunity-

effect. The unemployed state is also influenced by beliefs about the jobs that he or she can find. Decision and search theory would predict that a high-paying job is a job, particularly a high-paying job, is more likely to be found in a job search than those with lower-paying jobs. These beliefs are partly influenced by social contacts or neighbours, we have an effect. One example of belief-based effects is the *worker effect*, that is, that a high unemployment rate for individuals looking for work reduces the probability of finding a job (e.g., Schweitzer and Smith 1994). Interaction occurs when other individuals influence the focal individual. One reason for the effect in middle-class families from many studies, for example, was the influence of social contacts: 'the very presence of a high unemployment rate in the neighbourhood helps to keep the unemployed role models that help keep the unemployed state meaningful, that steady employment is the norm, and that family stability is the norm' (Schweitzer and Smith 1994: 56). In both the discouraged-unemployment among others influence and that his or her chances of leaving the unemployed state are altered.

that desire-based interactions are influenced by the existence of the social norm to earn one's income. Being unemployed and not live up to this norm, and this leads to embarrassment (Elster 1989a). This contributes to deviations from what is expected by the individual's reference groups. Since reference groups vary among individuals, the effect is not likely to be felt with equal

intensity by everyone. In particular, the more common it is to be unemployed in a group, the weaker the normative pressure is likely to be, and the less likely it is that an unemployed individual will experience such emotions. Being the only unemployed individual, furthermore, is likely to be a rather lonelier and duller existence than one in which many of one's friends and acquaintances also are unemployed. Thus, for several different reasons, one can expect that an increase in unemployment among an individual's friends and acquaintances will reduce the social and psychological 'costs' of being unemployed.

Clark (2003) presents some evidence suggesting that the unemployment of others indeed influences an individual's unemployment experience. Based on data from the British Panel Household Study, he found that it was easier for individuals to cope with unemployment (as measured by an index of subjective well-being) if they lived in places where many people were unemployed, or if others in the household also were unemployed. He also found that those whose subjective well-being fell the most on entering unemployment were more likely to search for new jobs and were less likely to remain unemployed later (see also Clark and Oswald 1994).

Clark's findings about the effects of the psychological and social costs of unemployment are parallel to those found in studies of the economic consequences of unemployment. The weight of this evidence suggests that increased unemployment benefits cause longer periods of unemployment. The main reason for this seems to be that higher benefits allow the unemployed to be more discriminating with regard to which jobs they accept and it allows them to somewhat reduce the effort they invest in searching for new jobs (see Holmlund 1998 for an overview). It seems likely that the social and psychological costs of being unemployed will have at least as strong an effect on the actions of the unemployed as these purely economic ones. The reason for this is that the variation in the non-pecuniary aspects is likely to be greater than the variation in the pecuniary ones and, as will be discussed below, the non-pecuniary consequences are likely to be self-reinforcing.⁴

To the extent that the unemployment of others influences an unemployed individual's subjective well-being and this, in turn, influences the unemployed individual's behaviour such that his or her chances of leaving the unemployed state are altered, we have an example of a desire-based interaction effect. As part of the project upon which this research

⁴ They are self-reinforcing in the same sense as a system of unemployment benefits that automatically became more generous when the unemployment level increased would be self-reinforcing. In both cases an increase in unemployment would set in motion processes that would generate more unemployment.

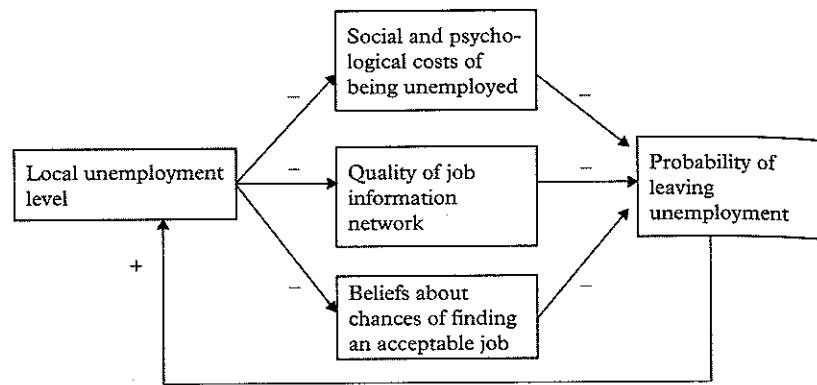


Figure 6.2. Unemployment as an endogenous process.

is based, we conducted a series of detailed interviews with unemployed youth in the Stockholm region. In these interviews, the importance of the unemployment of others was a recurrent theme, and one can find several examples that seem to indicate the importance of desire-based interactions. One interviewee said: 'If your friends are unemployed you do not think it is so bad to be unemployed since everyone else is. But if you are the only unemployed, you feel like an outsider.' And, he continued: 'If you do not have any unemployed friends, you don't have anything to do during the days. Then you would become restless and put more effort into finding a job.'

Figure 6.2 summarizes some of the discussion. An increase in the local unemployment level is likely to reduce the social and psychological costs of being unemployed (desire-based interactions), reduce the quality of the job information network (opportunity-based interactions) and reduce expectations about potential jobs (belief-based interactions). All these changes are likely to influence the unemployed individuals' behaviour in such a way that the probability of their leaving unemployment decreases, and this means that the local unemployment level will increase, everything else being the same.

If these types of social interaction effects are operating, one can expect endogenous processes to be important for changes in aggregate unemployment. A defining characteristic of an endogenous process is that the number of individuals who act in a certain way at a certain point in time in itself partly explains how many will adapt their behaviour at a later point in time. An exogenous event leading to a certain number of individuals becoming unemployed, then, can lead to many more individuals eventually becoming unemployed.

The disc
suspect t
portant r
ment. Bu
question.

The data
who ever
For thes
tion such
administ
at the en
employee
ment sp
88,000 i
range ha
were 20

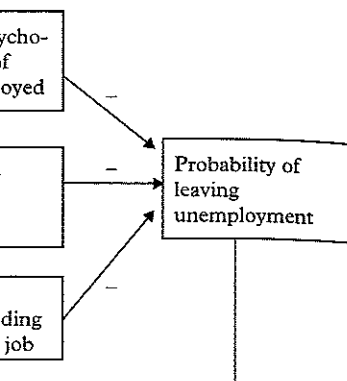
The r
that we
variables
public t
Stockho

⁵ As far as
interacti
to be o
Luttm
use; Be
pledges
rates;
explai
ing var

⁶ We her
county
county

⁷ The St
serve a
contain

⁸ The u
focus
employ



endogenous process.

ed interviews with unemployed
e interviews, the importance of
urrent theme, and one can find
the importance of desire-based
ur friends are unemployed you
ed since everyone else is. But if
like an outsider.' And, he con-
employed friends, you don't have
would become restless and put

cussion. An increase in the local
e social and psychological costs
ructions), reduce the quality of
-based interactions) and reduce
f-based interactions). All these
employed individuals' behaviour
their leaving unemployment de-
employment level will increase,

ts are operating, one can expect
for changes in aggregate un-
f an endogenous process is that
ertain way at a certain point in
will adapt their behaviour at a
leading to a certain number of
, can lead to many more indi-
d.

Transitions out of unemployment: statistical estimates

The discussion above thus suggests that there are good reasons to suspect that social interactions and endogenous processes play an important role in explaining temporal and spatial variations in unemployment. But whether or not they actually are important is still an open question. Answering this question is what we now seek to do.⁵

Data

The data set that we use contains information on all 20- to 24-year-olds who ever lived in the Stockholm metropolitan area during 1993–99.⁶ For these individuals we have traditional socio-demographic information such as age, sex, education and ethnicity (obtained from various administrative registers). We know in what neighbourhood they resided at the end of each calendar year,⁷ and for those who were ever unemployed we know the dates and exact lengths of all their unemployment spells.⁸ During the period January 1993–December 1999 about 88,000 individuals out of a total of about 226,000 individuals in that age range had at least one unemployment spell during the period when they were 20 to 24 years old.

The reason for restricting the analysis to a single metropolitan area is that we wish to hold constant one of the most important contextual variables: the tightness of the local labour market. Given the excellent public transportation system in this area, for all practical purposes the Stockholm metropolitan area can be viewed as one and the same labour

⁵ As far as we are aware, this is the first serious attempt to assess the importance of social interactions for unemployment durations. However, social interactions have been shown to be of importance for explaining other types of outcomes. See, for instance, Bertrand, Luttmer and Mullainathan (2000) and Mood (2004) for their role in explaining welfare use; Bearman and Brückner (2001) for their role in explaining the spread of virginity pledges; Glaeser, Sacerdote and Sheinkman (1996) for their role in explaining crime rates; Hedström (1994) and Hedström, Sandell and Stern (2000) for their role in explaining the diffusion of social movements; and Åberg (2003) for their role in explaining various demographic events.

⁶ We here define the 'Stockholm metropolitan area' as consisting of the entire Stockholm county, except for the following municipalities, which are situated at the outskirts of the county: Norrtälje, Sigtuna, Upplands Bro, Södertälje, Nykvarn and Nynäshamn.

⁷ The Stockholm metropolitan area is divided into 699 so-called SAMS areas, and these serve as our definition of neighbourhoods. The SAMS areas have been defined so as to contain socially homogeneous residential areas.

⁸ The unemployment data has been obtained from the so-called Händel database. We focus on 'open' unemployment, which means that we do not count among the unemployed those engaged in labour market training programmes and the like.

market. Thus, by restricting the analysis to a single metropolitan area, we reduce the risk of mistaking environmental effects in the form of geographical variations in labour market conditions for interaction-based peer-group effects.

Following the tradition of Hägerstrand (1967), we will assume that the structure of social interaction in part reflects actors' spatial locations: the closer two actors are to one another, the more likely they are to be aware of and influence each other's behaviour. The spatial distribution of a population for these reasons is likely to influence the web of social ties linking actors to one another and thereby also the outcome of the interaction-based process being analyzed (see also Hedström 1994).

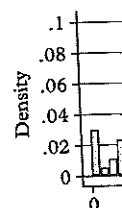
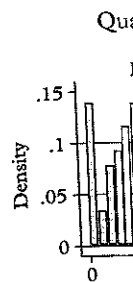
The reason for restricting the analysis to 20 to 24 year olds is that their significant others are to a large extent likely to be located in close geographical proximity. The processes we focus on are likely to be important for adults as well, but we then would have needed detailed information on the actual social networks linking the individuals to one another.

Neighbourhood variations

If social interactions are important then we should expect endogenous processes to generate differences in unemployment levels also between groups of interacting individuals who are similar to one another in terms of their labour-market-relevant characteristics. In order to examine whether or not this is the case, we examine the extent to which unemployment levels vary among neighbourhoods that are similar to one another in terms of their unemployment-relevant characteristics.

In order to identify neighbourhoods that resemble one another in terms of their unemployment-relevant characteristics, we estimated eighty-four logistic regression models, one for each month. We included only neighbourhoods with at least ten individuals in this age range. In the regression models the dependent variable indicated whether or not an individual was unemployed on the 15th of the month, and the independent variables measured the individual's age, sex, education, marital status, number of children, country of birth, whether or not (s)he was a student, and whether or not (s)he was a recent immigrant.⁹ Using these

⁹ We used sets of dummy variables to distinguish between the following educational levels: primary school only, vocational training school, high school degree and college degree; the following 'marital' statuses: living with parents, single household, and married or cohabiting; the following countries/regions of birth: Sweden, eastern Europe or former Soviet Union, Middle East or Africa, and the rest of the world. Being a 'recently' arrived



parameters
ity of be
probabil
at estim
observe
character
another
these ex

Figure
unempl
the resu
have an

immig
being
studen
The e
value

to a single metropolitan area, mental effects in the form of conditions for interaction-

(1967), we will assume that reflects actors' spatial locations: the more likely they are to be in a certain neighbourhood, the more likely they are to be in that neighbourhood. The spatial distribution of actors influences the web of social relations, and thereby also the outcome of the process (see also Hedström 1994).

For 20 to 24 year olds is that their spatial distribution is likely to be located in close proximity to the focus on are likely to be in that neighbourhood. This would have needed detailed data on the spatial distribution linking the individuals to one

we should expect endogenous factors to influence employment levels also between neighbourhoods that are similar to one another in terms of their characteristics. In order to examine the extent to which unemployment levels vary among neighbourhoods that are similar to one another in terms of their characteristics.

Neighbourhoods that resemble one another in terms of their characteristics, we estimated the expected unemployment level for each month. We included the characteristics of individuals in this age range. In order to indicate whether or not an individual was in the neighbourhood of the month, and the individual's age, sex, education, marital status, whether or not (s)he was a recent immigrant.⁹ Using these

between the following educational levels: high school degree and college degree; single household, and married or divorced; Sweden, eastern Europe or former Soviet Union; and Being a 'recently' arrived

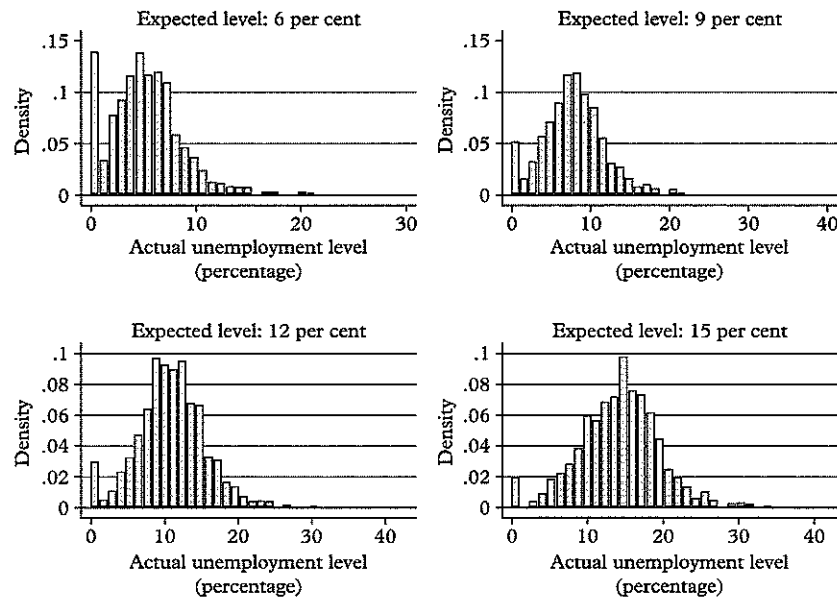


Figure 6.3. Variation in unemployment levels among neighbourhoods that are similar to one another in terms of their unemployment-relevant characteristics.

parameter estimates, we calculated each individual's predicted probability of being unemployed, and then we summarized these predicted probabilities for those within each neighbourhood. We thereby arrived at estimates of the unemployment level one would have expected to observe each month in a neighbourhood on the basis of the demographic characteristics of its members. Two neighbourhoods are similar to one another in their unemployment-relevant demographic characteristics if these expected unemployment levels are approximately the same.

Figure 6.3 compares four sets of neighbourhoods. In the first set the unemployment-relevant demographics were such that, on the basis of the results from the logistic analyses, one would have expected them to have an unemployment level of 6 per cent.¹⁰ In the second set one would

immigrant was defined as having arrived in Sweden during the previous three years, and being a 'student' was defined on the basis of whether or not the individual had received student allowance ('studiebidrag') during the year.

¹⁰ The expected levels are equal to the predicted levels rounded to the nearest integer value.

have expected an unemployment level of 9 per cent, and in the third and fourth sets 12 per cent and 15 per cent respectively.¹¹

The results in figure 6.3 clearly show that unemployment levels vary more between neighbourhoods than one would expect them to do on the basis of their unemployment-relevant demographics. In approximately 50 per cent of these cases the actual unemployment level deviated by more than 25 per cent from the expected level.

As noted above, a likely reason for these 'excessive' differences between neighbourhoods is the existence of social interaction effects that set in motion endogenous social processes within certain neighbourhoods. But before we can endorse such an interpretation we need to examine whether the social interaction effect indeed is sufficiently strong to generate such a pattern.

Social interaction effects

In this subsection we seek to assess the extent to which individuals' unemployment-relevant actions are influenced by their peers. The ideas that have guided our analysis are displayed in the Coleman-like figure 6.4. We know that the chance that an unemployed individual will escape unemployment is influenced by his or her actions, for example, the extent and intensity of the individual's job search. We furthermore know that these actions vary among individuals with different attributes such as age, sex, education and ethnicity. As detailed above, there also are strong reasons to suspect that these actions are influenced by the unemployment level among their peers. Obviously their chances of leaving unemployment are not only due to their own actions but are also influenced by the tightness of the labour market (for example, by the number of vacant jobs in relation to the number of unemployed individuals) and by their attributes (reflecting employers' preferences for hiring individuals with certain characteristics). Finally, changes in the individual's probabilities of leaving unemployment will influence the unemployment level at the next point in time. However, this final stage of the analysis, which concerns the transition from the individual to the social, is not part of the statistical analysis. For that purpose the agent-based model will be used.

Since the process we analyze unfolds over time, and since the outcome variable we are interested in – leaving unemployment – is a discrete

¹¹ These four sets represented 29 per cent of all monthly neighbourhood observations and they appear representative of the others.

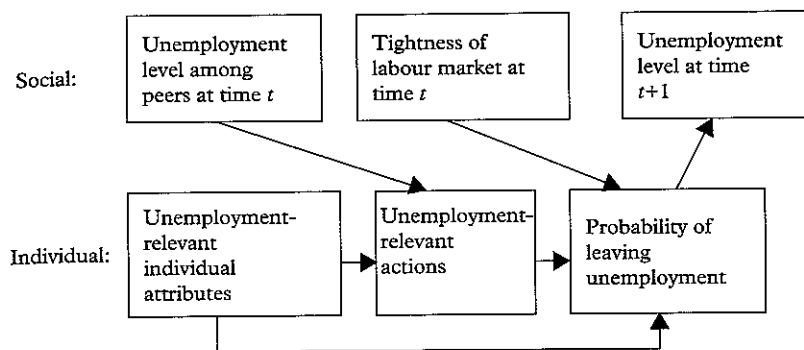


Figure 6.4. Social and individual components of the outflow from unemployment.

event, the statistical model we use is a so-called discrete-time-event history model (see Allison 1982). This essentially is a regular logistic regression model where the unit of analysis has been changed from persons to 'person weeks'. That is to say, before the parameters of the logistic model are estimated, the data is changed in such a way that each person contributes as many observations as the number of weeks that (s)he was at risk of leaving unemployment. An individual who was unemployed for two weeks thus contributes only two observations, while an individual who was unemployed for fifty weeks contributes fifty observations. The set of 87,924 individuals included in the analysis contributed a total of 2,463,079 person weeks.

Unfortunately, our data set does not include any information about what the unemployed individuals did to affect their chances of leaving unemployment. Therefore, we must estimate the parameters of a so-called reduced form model which directly relates an unemployed individual's probability of leaving unemployment to the tightness of the labour market, the attributes of the individual in question, and the unemployment level among his or her peers.

The first model in table 6.1 relates an individual's probability of leaving unemployment during a specific week to the unemployment level among his or her neighbourhood peers. The unemployment level among peers is calculated as the proportion of unemployed 20- to 24-year-olds in the neighbourhood at the end of the week preceding the week being analyzed. The logistic regression coefficient associated with this variable is less than zero, which means that the social interaction effect is in the expected direction. It suggests that the higher the unemployment level is among an unemployed individual's peers, the lower the likelihood is of

Table 6.1. *Logistic regression model of the probability of leaving unemployment: regression coefficients, with z statistics in parentheses*

	Model 1	Model 2
Unemployment level among peers (at the end of the preceding week)	-4.086 (82.99)	-2.087 (-33.59)
Woman		0.132 (25.98)
Age		-0.023 (-9.74)
Vocational training		0.027 (3.83)
High school education		0.137 (20.16)
College education		0.206 (21.71)
Immigrant from eastern Europe or former Soviet Union		-0.138 (-8.40)
Immigrant from Middle East or Africa		-0.192 (-18.84)
Immigrant from the rest of the world		-0.014 (-1.53)
Less than 3 years in Sweden		-0.455 (-27.64)
3-5 years in Sweden		-0.044 (-3.28)
Married		-0.034 (-2.76)
No. of children		-0.055 (-6.16)
Previous unemployment experiences (no. of weeks/10)		-0.019 (-19.12)
Number of unemployed per vacant job (at the beginning of the month)/100		-0.034 (-0.24)
Length of current unemployment spell (no. of weeks)/10		0.319 (65.40)
Square of the length of current unemployment spell		-0.045 (-51.26)
Constant	-2.085 (363.35)	-2.145 (33.69)
Annual and monthly dummy variables included	No	Yes
Log likelihood	-644312.97	-627468.57

him or her leaving unemployment. The value of -4.086 suggests a substantial social interaction 'effect'. To make an out-of-sample prediction, it suggests that, if everyone in the peer group were unemployed, the individual's probability of leaving unemployment would be only about 1.7 per cent of what it would have been had no one been unemployed.

Obviously, much of this so-called social interaction effect is likely to be due to individual heterogeneity across neighbourhoods, which we must control for. If we did not, we would seriously overestimate the extent to which individuals are influenced by others. In the second model, we therefore include variables to control for relevant individual-level differences: sex, age, education (highest degree), country of

birth, number of children, number of unemployed, previous unemployment, control for age, to influence (but it is not). We also control for years of yearly unemployment, number of children, and more labour market experience. The current duration of unemployment.

The number of children among the unemployed, leaving unemployment, labour market experience suggests that individual characteristics per cent of what it would have been. But, on the other hand, be treated as a control variable.

The variables they are included in the same model, the same variables with age, unemployment, category of unemployment, from the difficulty of finding a job, that in the chance of finding a job, the work arrived at that more than six months or her own individual

Leaving unemployment: regression

Model 1	Model 2
86 (82.99)	-2.087 (-33.59)
	0.132 (25.98)
	-0.023 (-9.74)
	0.027 (3.83)
	0.137 (20.16)
	0.206 (21.71)
	-0.138 (-8.40)
	-0.192 (-18.84)
	-0.014 (-1.53)
	-0.455 (-27.64)
	-0.044 (-3.28)
	-0.034 (-2.76)
	-0.055 (-6.16)
	-0.019 (-19.12)
	-0.034 (-0.24)
	0.319 (65.40)
	-0.045 (-51.26)
85 (363.35)	-2.145 (33.69)
Yes	
312.97	-627468.57

The value of -4.086 suggests
To make an out-of-sample
in the peer group were un-
leaving unemployment would
d have been had no one been

interaction effect is likely to
neighbourhoods, which we
d seriously overestimate the
ed by others. In the second
control for relevant individ-
(highest degree), country of

birth, number of years residing in Sweden, marital status, number of children and previous unemployment experiences measured as the total number of weeks the individual had been unemployed before the current unemployment period started. The variable measuring the extent of previous unemployment experiences has been included in order to control for unobserved and otherwise uncontrolled heterogeneity likely to influence the probability of an individual leaving unemployment (but it also may pick up so-called scarring effects of unemployment). We also control for the tightness of the labour market by including a set of yearly and monthly dummy variables and a variable measuring the number of unemployed per job vacancy at the beginning of each calendar month in the Stockholm county (according to statistics from the labour market authorities). The variable measuring the length of the current unemployment spell has been included to control for so-called duration dependence.

The most important result in model 2 is that the unemployment level among neighbourhood peers has a substantial effect on the probability of leaving unemployment even after we control for all these individual and labour market attributes. The logistic regression coefficient of -2.087 suggests that, if all the neighbourhood peers were unemployed, the individual's risk of leaving unemployment would be only about 12.4 per cent of what it would have been had no one been unemployed. But, once again, this is an out-of-sample prediction and should therefore be treated with some caution.

The effects of some of the other covariates are also interesting, but they are not our primary concern in this chapter. The results for these variables may be briefly summarized as follows. Everything else being the same, they suggest that women are more likely to leave unemployment than men; that the probability of leaving unemployment decreases with age; that those with higher education have better chances of leaving unemployment than those with lower education (the omitted reference category is those with compulsory schooling or less); that immigrants from eastern Europe and from the former Soviet Union have a more difficult time leaving unemployment than people born in Sweden, and that immigrants from the Middle East and Africa have even lower chances of leaving unemployment (though immigrants from the rest of the world do not differ from those born in Sweden); that being a recently arrived immigrant reduces the possibilities of leaving unemployment; that married persons have a smaller chance of leaving unemployment than single persons; that the more children a person has, the lower his or her chances of leaving unemployment are; and finally, that the more an individual has been unemployed in the past, the lower his or her

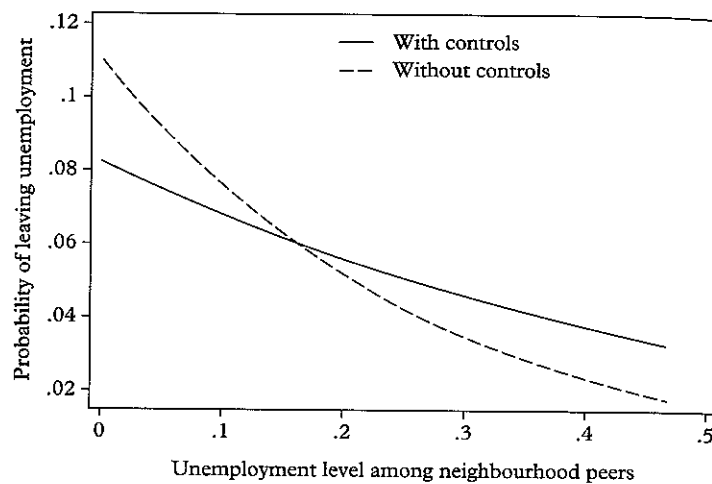


Figure 6.5. Estimated strength of social interaction effects for an average person.

chances of leaving the current unemployment spell are. The coefficients associated with the variables measuring the length of the current unemployment spell suggest that the probability of leaving unemployment gradually increases and reaches its peak when the individual has been unemployed for thirty-six weeks.

For an average individual these results imply the social interaction effects described in figure 6.5. The graphs show that the probability of an unemployed individual leaving unemployment is considerably influenced by the unemployment level among neighbourhood peers, also when all the covariates of table 6.1 are controlled.

Selection effects that we have not been able to control for may have led to an upward bias in these estimates. But our crude measure of the reference group variable is likely to have led to a bias in the opposite direction. It is unclear how these two sources of error jointly influence our estimates, but they should at least partly cancel one another out. A lower bound on the social interaction effect can be arrived at by a fixed-effect specification that controls for all time-invariant differences between the neighbourhoods. Using such a technique undoubtedly means that one introduces excessive controls and therefore biases the estimate downwards, but analyses not reported here show that, even with such excessive controls, with a 95 per cent confidence level the true logistic regression coefficient measuring the strength of the social interaction

effect is to be fo
as large as -2.3

All in all, the
among neighbo
ability that a
Although some
differences bet
control for, it
these rather su
sition from the
examine the in
likely to be ob
based model.

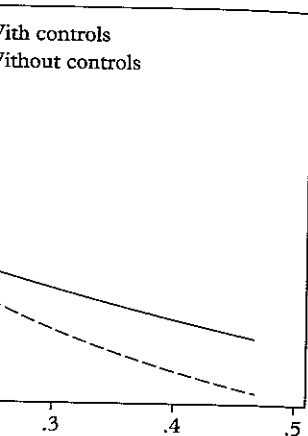
A sim

As mentioned
use empiricall
the social-level
order to conv
models can be
ual-level rese
will be incorp
quantitative
micro-macro
figure 6.1).

In order to
following ma
stylized agen
sion results o
which the ag
outcomes th
after we dev
assumptions
from the em

¹² This estimat
total 121,72
the fixed-eff

¹³ Results not
fixed-effect
the individua



ing neighbourhood peers

of social interaction effects for an

ment spell are. The coefficients
the length of the current un-
employment of leaving unemployment
when the individual has been

ts imply the social interaction
s show that the probability of an
employment is considerably influ-
ing neighbourhood peers, also
controlled.

able to control for may have led
But our crude measure of the
e led to a bias in the opposite
sources of error jointly influence
rtly cancel one another out. A
ect can be arrived at by a fixed-
time-invariant differences be-
technique undoubtedly means
d therefore biases the estimate
ere show that, even with such
confidence level the true logistic
length of the social interaction

effect is to be found in the interval -3.24 to -1.50 , with a point estimate as large as -2.37 .¹²

All in all, these results strongly suggest that the unemployment level among neighbourhood peers has a considerable influence on the probability that an unemployed individual will leave unemployment. Although some of these social interaction effects are likely to be due to differences between the individuals that we have not been able to control for, it seems highly unlikely that such factors could wipe out these rather substantial peer-group effects.¹³ In order to make the transition from the level of the individual to the level of the social, and to examine the implications of these results for the unemployment levels likely to be observed, we must incorporate the results into an agent-based model.

A simple agent-based model of unemployment

As mentioned above, the core idea of the approach advocated here is to use empirically calibrated agent-based models (ECA models) to derive the social-level implications of a set of quantitative research results. In order to convey what type of model we have in mind and how such models can be used for assessing the social outcomes implied by individual-level research findings, the empirical results of the previous section will be incorporated into an agent-based model. If this approach is used, quantitative research comes to have a direct bearing on the so-called micro-macro link discussed by Coleman and others (see arrow 3 in figure 6.1).

In order to illustrate the logic of the approach, we will proceed in the following manner. First, we inject some realism into the type of highly stylized agent-based model used in chapter 4. We use the logistic regression results of table 6.1 to arrive at a more plausible model of the ways in which the agents influence one another, and then examine the social outcomes they bring about under these more realistic conditions. Thereafter we develop the ECA model by replacing many of the simplified assumptions of this stylized agent-based model with information derived from the empirical analysis. The ECA model will be used as a virtual

¹² This estimate is based on a 5 per cent random sample of the unemployment spells, in total 121,727 person-weeks. For computational reasons it was not feasible to estimate the fixed-effect model on the total population.

¹³ Results not reported here show that this conclusion remains the same when so-called fixed-effect specifications are used to control for all time-invariant differences between the individuals' neighbourhoods.

laboratory to examine how social outcomes are likely to be affected by various changes at the level of the individual.

The agent-based models of chapter 4 were used to analyze how social interactions among actors were likely to bring about changes in the actors' beliefs and desires, and thereby also in their actions. In those models it was assumed that actors' beliefs/desires changed if and only if a majority of their neighbours had beliefs/desires that were different from their own. This type of agent-based model can be made more realistic by implementing an evidence-based action rule. If we assume, as was done in chapter 4, that the actors' opportunities are such that they can be in only two states, the first regression model of table 6.1 says that the probability that an actor will change state/action at a specific point in time is given by the expression:

$$p_{jt} = \frac{1}{1 + e^{2.085 + 4.086 \times U_{jt-1}}}$$

where U_{jt-1} equals the proportion of the neighbours who were in the same state or acted in the same way as the focal actor at the previous point in time.¹⁴ The equation says that the larger the proportion of the neighbours that acted in the same way as the focal actor, the less likely it was that the actor would change action.

In order to examine the social patterns that emerge when agents' actions are decided on the basis of this rule, we proceed in the same manner as in chapter 4. We assume that 2,500 actors are placed on a lattice (torus) with fifty rows and fifty columns. We start with an entirely random action pattern and then we examine the social patterns that emerge when the agents interact and influence one another. One important difference between these analyses and those in chapter 4 is that we now focus on the actions as such and not on the underlying beliefs and desires of the actors. It would have been desirable also to include beliefs and desires in the analysis, but we do not have any empirical information about them.

A typical initial action pattern looks like the upper-left graph of figure 6.6. Black areas identify actors who acted in one way (call their action a B-action), and white areas identify those who acted in the other way (call their action a W-action). In the simulation reported in figure 6.6, 40 per cent of the actors performed a B-action and 60 per cent performed a W-action at the outset of the analysis.

¹⁴ The results of these analyses can be interpreted as either referring to the states in which the actors are or in terms of their actions. To simplify the presentation, hereafter the results are presented in action terms.

es are likely to be affected by
al.

ere used to analyze how social
bring about changes in the
also in their actions. In those
desires changed if and only if a
sires that were different from
can be made more realistic by
le. If we assume, as was done
s are such that they can be in
el of table 6.1 says that the
/action at a specific point in

$6 \times U_{t-1}$

neighbours who were in the
e focal actor at the previous
e larger the proportion of the
e focal actor, the less likely it

s that emerge when agents'
ule, we proceed in the same
2,500 actors are placed on a
nns. We start with an entirely
hine the social patterns that
nce one another. One import-
those in chapter 4 is that we
on the underlying beliefs and
sirable also to include beliefs
ve any empirical information

the upper-left graph of figure
n one way (call their action a
o acted in the other way (call
reported in figure 6.6, 40 per
and 60 per cent performed a

either referring to the states in which
bly the presentation, hereafter the

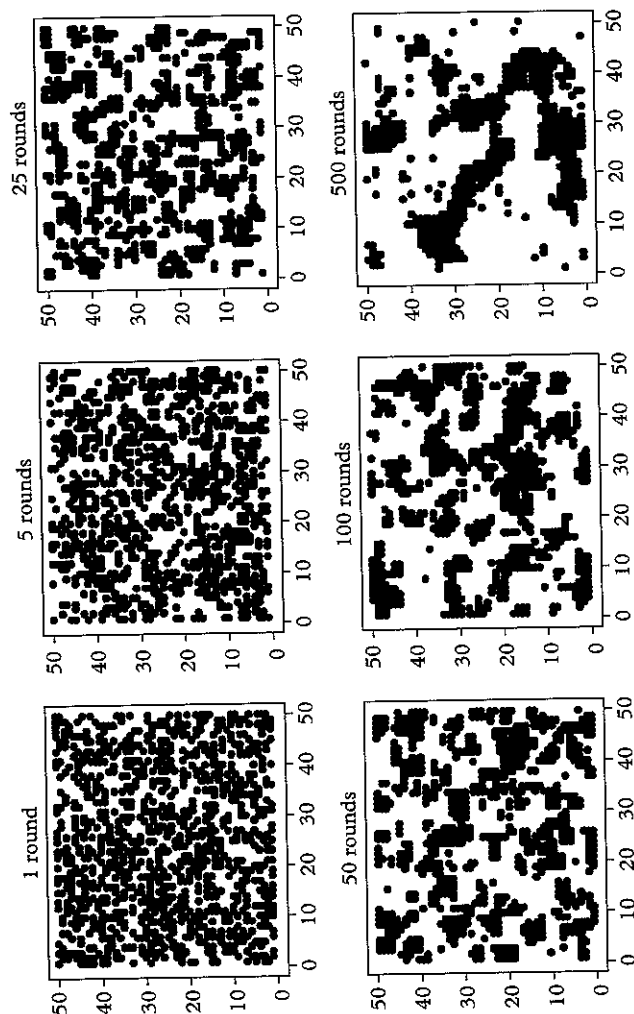


Figure 6.6. Typical action patterns in a population of 2,500 actors who socially interact with four neighbours on the basis of an empirically calibrated action rule.

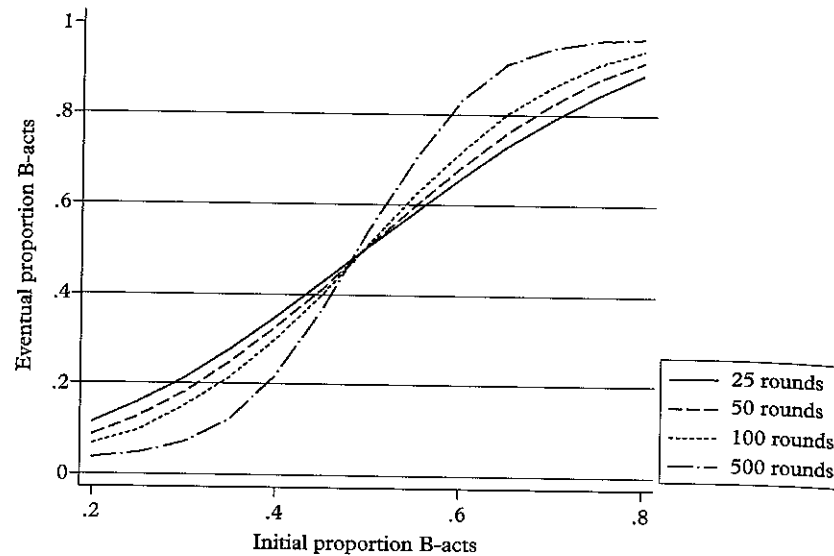


Figure 6.7. Summary of the results of 5,200 agent-based analyses in which 2,500 agents interact on a lattice (torus) with 50 rows and 50 columns on the basis of an empirically calibrated action rule.

We take social interactions into account by assuming that the actors are directly influenced by their four immediate neighbours (see Figure 4.2). These neighbours influence the focal actor's probability of changing his or her action in the manner described by the logistic regression equation. The social patterns that emerge under these conditions typically look like those in figure 6.6. Although we start with an entirely unstructured social pattern, a highly segregated pattern emerges rather quickly. Already when the actors have interacted and influenced one another over five rounds, segregated patterns start to emerge. As the interaction process proceeds, the extent of clustering and segregation increases. Thus it seems that social interaction processes can bring about highly segregated social patterns also when the agents act on the basis of plausible assumptions about the strength of social interaction.¹⁵

Figure 6.7 summarizes the results of a large number of agent-based analyses like these, and gives some additional insights into the social

¹⁵ Bruch and Mare (2004) found that the use of plausible probabilistic decision rules in a traditional Schelling (1971) model of residential segregation did not generate the highly segregated patterns that Schelling models normally generate. They discussed whether their finding could be generalized to social-influence processes more generally. These results suggest that they cannot.

outcomes that the clustering effect. The processes also have in common, the social pattern and if the action is common. Figure 6.7 shows the agents' interaction is likely to be.

On a very abstract level, the similarities with the social pattern are highly dependent on the attributes of the interactions may be in unemployment, therefore we must

Although the interaction is into the agent-based reality, the remaining one can rightfully say that the intended purposes approximate the social results. On the other hand, the assumed chance balance to real-world based on such simulation outcomes implies that these analyses do not. If we do not allow for level predictions. Finally, state-of-the-art. Most of the time, of complex history.

Any particular event is a complicated cause and effect, short and simple, they are infinite.

Faced with a long length, in practice, their most recent start from a previous phenomena as

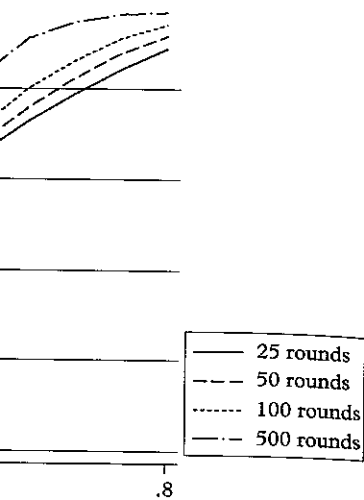


Figure 6.6: Results of 5,200 agent-based analyses in a lattice (torus) with 50 rows and 50 columns, calibrated to a specific action rule.

ent by assuming that the actors interact with their immediate neighbours (see Figure 6.6). The focal actor's probability of interacting with a neighbour is described by the logistic function. As the number of rounds increases, the patterns start to emerge. As the number of clustering and segregation processes can bring about changes in the agents act on the basis of the results of social interaction.¹⁵

A large number of agent-based analyses have provided additional insights into the social

possible probabilistic decision rules in a world where segregation did not generate the highly segregated patterns. They discussed whether the processes are more generally. These

outcomes that these types of processes tend to generate. In addition to the clustering effects shown in figure 6.6, it seems that these types of processes also have important magnifying effects. If an action is common, the social interaction process makes it even more common, and if the action is uncommon, the interaction process makes it even less common. Figure 6.7 also shows that the longer or the more frequently the agents interact with one another, the stronger this magnifying effect is likely to be.

On a very abstract level, the patterns in figures 6.6 and 6.7 show some similarities with the unemployment patterns found above. Both types of pattern are highly segregated and they cannot be explained by reference to the attributes of the actors. Such similarities suggest that social interactions may have been important in generating the spatial variation in unemployment, but obviously the evidence is far from conclusive, and therefore we must carry out more detailed analyses.

Although the inclusion of the results of the logistic regression analysis into the agent-based model have reduced the gap between model and reality, the remaining gap is considerable by any measure. For this reason one can rightfully wonder whether or not this type of model can serve the intended purpose of being the micro-macro link that allows us to approximate the social outcomes implied by this set of micro-level statistical results. One obvious discrepancy between the model and reality is the assumed checkerboard structure, which shows little or no resemblance to real-world social structures. It is far from certain that a model based on such simplifying assumptions can accurately generate the social outcomes implied by the micro-level findings. Similarly, the agents of these analyses do not have much in common with real-world individuals. If we do not allow for real-world heterogeneity, it is likely that the social-level predictions derived from the model will be systematically biased. Finally, state-of-nature models such as these are always a little problematic. Most of the social phenomena that we seek to explain are the results of complex historical processes. As David Lewis once put it:

Any particular event that we might wish to explain stands at the end of a long and complicated causal history. We might imagine a world where causal histories are short and simple; but in the world as we know it, the only question is whether they are infinite or merely enormous. (Lewis 1983: 214)

Faced with a world consisting of causal histories of nearly infinite length, in practice we can hope only to provide reliable information on their most recent history. Instead of basing the analysis on models that start from a presocial random state, it seems safer to take certain social phenomena as given and incorporate them into the agent-based model.

The realism of the model is thereby enhanced, which gives us more faith in the results derived from it.

For all these reasons, highly stylized agent-based models are not likely to give a good approximation of the social outcomes implied by a set of statistical results. The model to be used must have more empirical texture than these models have in order to be useful for this purpose. Our strategy for arriving at such a model can be described as follows:

- 1 Hypothetical agents should be replaced with virtualized replicas of heterogeneous real agents.
- 2 The checkerboard structure should be replaced with real spatial or social structures.
- 3 The structure as well as the strength of social interaction should be estimated with real data.
- 4 Important real-world events known to influence the outcome to be explained should be incorporated into the model.

But – and this is at the heart of our approach – the logic of the analysis should remain the same as in traditional agent-based analyses. That is to say, it is the actions of and interactions between the agents that should generate the social patterns that emerge, and by altering various aspects of the simulation setup one ascertains what effects these changes may have on the outcomes.

An empirically calibrated agent-based model of unemployment

In order to construct an empirically calibrated agent-based (ECA) model of unemployment, instead of basing the analyses on 2,500 hypothetical actors we should use virtual replicas of the individuals who actually experienced unemployment during this period as our agents. Instead of assigning them positions on a checkerboard-like structure, we should assume that they resided where their real-world counterparts actually did and that they interacted with virtual replicas of their actual neighbourhood peers. And instead of just making assumptions about how agents' actions are influenced by the actions of others, we should use the results from the large-scale data analyses presented above to empirically specify what the functional relationships look like.

The agents of the ECA model thus are virtualized replicas of all the 20 to 24 year olds in the Stockholm metropolitan area who were ever unemployed during the period January 1993–December 1999. All in all, 87,924 agents are included in the analyses. These agents retain the true social and demographic characteristics of their real-world counterparts

(see table 6.1 for count). We also u
experience any u
unemployed in th
only those who e
with their true nei
ence one another

In the analyses
counterparts actu
just as their real-
analysis if their r
ropolitan area, if
than 300 days (w
this age range co

Their probabili
of factors: (1) th
unemployment
tightness of the l
their probability
logistic regressio

The simulation
agents leave un
the number of
to introduce ch
the agents' exit
the expected n
act and influen
plier effect on
because a chan
level of unemp
exit probabiliti
level in each r
probabilities o
unemployment
changes in exit
from January

Figure 6.8
this period an

¹⁶ In order to hi
removed from
moving avera
horizontal axe

anced, which gives us more faith
agent-based models are not likely
cial outcomes implied by a set of
used must have more empirical
er to be useful for this purpose.
el can be described as follows:

anced with virtualized replicas of
be replaced with real spatial or
of social interaction should be
to influence the outcome to be
o the model.

proach – the logic of the analysis
agent-based analyses. That is to
between the agents that should
, and by altering various aspects
what effects these changes may

agent-based model

calibrated agent-based (ECA)
ing the analyses on 2,500 hypo-
replicas of the individuals who
uring this period as our agents.
a checkerboard-like structure,
re their real-world counterparts
h virtual replicas of their actual
ust making assumptions about
he actions of others, we should
ta analyses presented above to
relationships look like.

re virtualized replicas of all the
etropolitan area who were ever
1993–December 1999. All in all,
es. These agents retain the true
of their real-world counterparts

(see table 6.1 for a description of the characteristics we take into account). We also use information on the 20 to 24 year olds who did not experience any unemployment when we calculate the proportion of unemployed in the neighbourhoods, but the agents in our analyses are only those who experienced any unemployment. The agents interact with their true neighbourhood peers, and the extent to which they influence one another is given by the results of the empirical analysis.

In the analyses the agents become unemployed when their real-world counterparts actually became unemployed. They age, move and so forth just as their real-world counterparts did. The agents also exit from the analysis if their real-world counterparts move from the Stockholm metropolitan area, if they turn 25, or if they have been unemployed for more than 300 days (which is the maximum number of days that individuals in this age range could receive unemployment benefits).

Their probability of leaving unemployment is influenced by three sets of factors: (1) their own social and demographic characteristics, (2) the unemployment level among their neighbourhood peers, and (3) the tightness of the labour market. The ways in which these factors influence their probability of leaving unemployment are given by the second logistic regression equation in table 6.1.

The simulation model focuses on how changes in the rate at which the agents leave unemployment influence the level and spatial variation in the number of unemployed. The idea behind the virtual experiments is to introduce changes in the extent to which different factors influence the agents' exit probabilities. Such changes will have a direct effect on the expected number of unemployed agents but, since the agents interact and influence one another, it will also have an indirect social-multiplier effect on the unemployment level. The agents are interdependent because a change in the exit probability of some agents will change the level of unemployment in their neighbourhoods, and this will change the exit probabilities of others. At the end of each week the unemployment level in each neighbourhood is updated and allowed to influence exit probabilities during the following week. This will in turn influence unemployment levels at the end of that week, which will lead to further changes in exit probabilities, and so on, throughout the 364-week period from January 1993 to December 1999.

Figure 6.8 describes how the unemployment level developed during this period and the outcomes of some of the virtual experiments.¹⁶ The

¹⁶ In order to highlight general trends and differences, all seasonal variations have been removed from the graphs in figures 6.9–6.11 with a smoothing routine. All graphs report moving averages based on the 26-week periods before and after each date on the horizontal axes.

unemployment level was very high during 1993–95 but it fell rapidly thereafter (see the dash-dot line and use the right-hand axis). At its peak more than one young person out of ten was unemployed, and at its lowest point about one young person out of twenty-five was unemployed. The ECA simulations have a counterfactual purpose. We use them to assess how the unemployment level, overall as well as in different neighbourhoods, is likely to have differed if the social interaction effects had been different from what they actually were.

The baseline model in these analyses is a simulation based on the actual parameter estimates found in the second model of table 6.1. This baseline model serves two purposes. First, it allows us to examine the extent to which the agent-based model can bring about the social outcomes it was intended to explain. Second, it serves as a point of reference for the virtual experiments. As far as the first of these purposes is concerned, the results suggest that the model is fairly successful. The correlation between the actual unemployment level in the various neighbourhoods at different points in time and the unemployment levels brought about when we assume that the agents' actions are governed by the baseline parameters is as high as 0.84.

To simplify the comparisons between the baseline simulation and the various virtual experiments, the overall unemployment level brought about each week under the baseline regime is set equal to 100, and the unemployment levels brought about by the experimental regimes are expressed as a percentage of the baseline level. The solid line in figure 6.8 is the baseline reference point, and the long-dashed line shows how the unemployment level would have changed if the social interaction effect (as measured by the logistic regression coefficient) was 50 per cent higher than it actually was but everything else remained the same. This increase in the extent to which the actors were influenced by others would have increased the number of unemployed by 8 per cent during an average week (from now on, use the left-hand axis). During the high unemployment period, the increase would have been as high as 10 per cent. It should be noted that these differences are entirely due to changes in the rate at which unemployed individuals leave unemployment. In both scenarios, the inflow of unemployed individuals is identical.

The medium-dashed line in figure 6.8 shows how the unemployment level would have changed if the social interaction effect was 50 per cent *lower* than it actually was (once again as measured by the size of the logistic regression coefficient). The results are similar, but in the opposite direction to those discussed in the previous paragraph. This change in

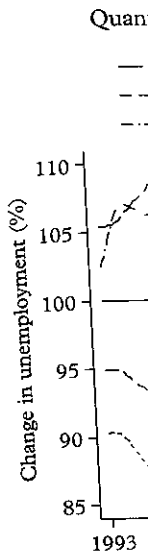


Fig
hol

the extent
ced the n
during the
during the

The sho
no social
unemploy
peers. On
about un
about in t
ment lev
what the
unemplo
according
effects.

In term
consider
would h
the high
interacti
about 5

ing 1993–95 but it fell rapidly the right-hand axis). At its peak was unemployed, and at its out of twenty-five was unemployed, and at its a counterfactual purpose. We ment level, overall as well as in ave differed if the social inter- that they actually were.

s is a simulation based on the second model of table 6.1. This first, it allows us to examine the can bring about the social out- , it serves as a point of reference the first of these purposes is model is fairly successful. The ment level in the various neigh- and the unemployment levels e agents' actions are governed 0.84.

the baseline simulation and the unemployment level brought game is set equal to 100, and t by the experimental regimes baseline level. The solid line in and the long-dashed line shows ve changed if the social inter- c regression coefficient) was 50 t everything else remained the h the actors were influenced by r of unemployed by 8 per cent use the left-hand axis). During ease would have been as high t these differences are entirely mployed individuals leave un- w of unemployed individuals is

shows how the unemployment interaction effect was 50 per cent s measured by the size of the s are similar, but in the oppos- ous paragraph. This change in

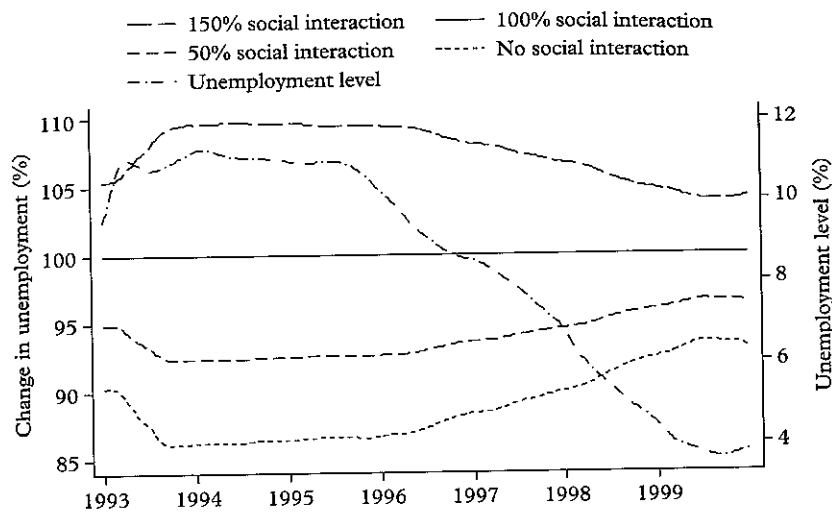


Figure 6.8. Actual and simulated unemployment levels in the Stockholm metropolitan area.

the extent to which the actors influence one another would have reduced the number of unemployed individuals by more than 5 per cent during the high employment period and by slightly less than 5 per cent during the latter half of the period.

The short-dashed line shows what would have happened if there were no social interaction effects at all, that is, if the probability of leaving unemployment were unaffected by the unemployment level among peers. Once again, the unemployment levels that the actors would bring about under these conditions differ considerably from those brought about in the baseline simulation. Under these conditions the unemployment levels would have been between 86 per cent and 93 per cent of what they were under the baseline set up. On average, the number of unemployed individuals would have been 89 per cent of what it was according to the baseline simulation had there been no social interaction effects.

In terms of economic as well as social costs, these differences are of considerable interest. These analyses suggest that on the average there would have been 990 fewer young people unemployed each week during the high unemployment years 1993–95 if there had not been any social interaction effect at all. This means that the social interaction generated about 51,000 additional unemployment weeks per year, which should be

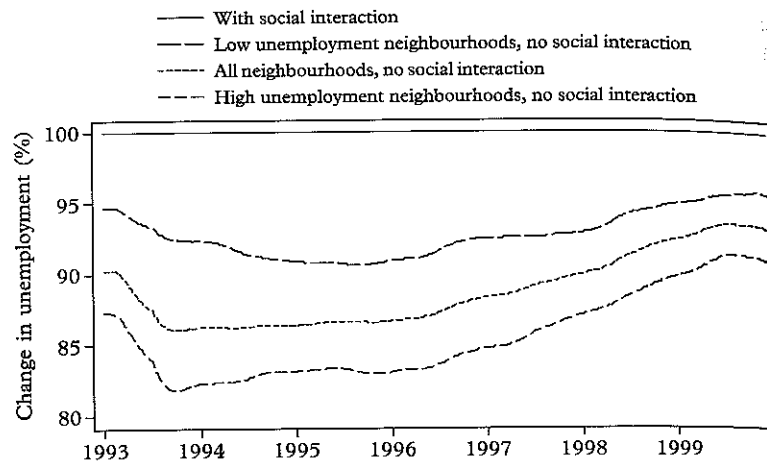
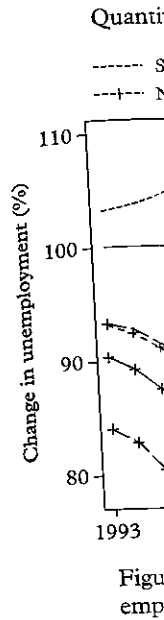


Figure 6.9. Unemployment levels and social interactions in low and high unemployment neighbourhoods in the Stockholm metropolitan area.

seen in relation to the fact that there were slightly fewer than 82,000 20 to 24 year olds who lived in the Stockholm metropolitan area during these years.

As can be seen from figure 6.9, social interactions are likely to influence not only the overall level of unemployment but also its spatial variation. Once again, to simplify the comparisons between the baseline and the various scenarios, the overall unemployment level brought about each week under the baseline regime is set equal to 100, and the unemployment levels brought about by the virtual experiments are expressed as a percentage of the baseline level. Figure 6.9 shows that social interactions tend to magnify the differences between low unemployment and high unemployment neighbourhoods.¹⁷ That is, if the unemployment level is higher in certain neighbourhoods than in others, perhaps because of demographic differences between the individuals residing in the neighbourhoods, social interactions are likely to magnify these differences since the multiplier effect will be greater in the high unemployment areas. While the number of unemployed in the low unemployment areas would have been about 93 per cent of what it

¹⁷ The 10 per cent neighbourhoods with the lowest average unemployment during 1993 were defined as 'low unemployment neighbourhoods' and the 10 per cent with the highest average unemployment as 'high unemployment neighbourhoods'.

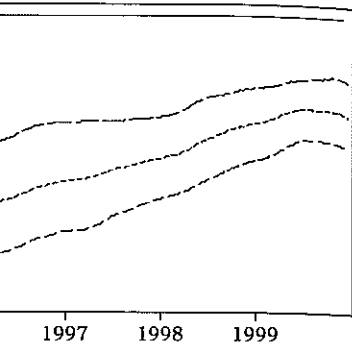


actually was
ing figure for

Since data
difficult to
compares w
unemploy
the social i
that the ag
regression
one another
about are
the figure
on differ
interaction
when the
dashed li
levels had
agents wo

¹⁸ The str
previous

neighbourhoods, no social interaction
social interaction
neighbourhoods, no social interaction



and social interactions in low and
neighbourhoods in the Stockholm metropolitan

were slightly fewer than 82,000 20
Stockholm metropolitan area during

cial interactions are likely to influ-
employment but also its spatial
comparisons between the baseline
employment level brought about
is set equal to 100, and the un-
the virtual experiments are ex-
line level. Figure 6.9 shows that
the differences between low un-
neighbourhoods.¹⁷ That is, if the
n neighbourhoods than in others,
ferences between the individuals
interactions are likely to magnify
effect will be greater in the high
number of unemployed in the low
en about 93 per cent of what it

west average unemployment during 1993
neighbourhoods' and the 10 per cent with the
employment neighbourhoods'.

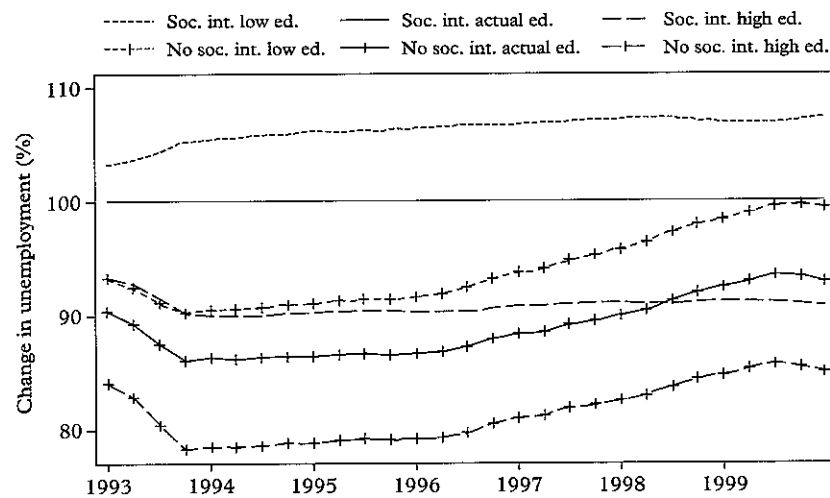


Figure 6.10. Effects of social interactions and education on the unemployment level in the Stockholm metropolitan area.

actually was had there been no social interaction effects, the corresponding figure for the high unemployment areas was 85 per cent.

Since data to examine social interaction effects are rarely available, it is difficult to have any intuitive sense of how the magnitude of these effects compares with the effects of factors usually considered in studies of unemployment. For this reason figure 6.10 compares the magnitude of the social interaction and the educational effects. As before, we assume that the agents base their actions on the results of the second logistic regression model in table 6.1, and we allow them to act and to influence one another week by week. The unemployment levels they then bring about are those shown in figure 6.10. The various outcomes shown in the figure are due to different experimental setups, that is, they are based on different assumptions about the strength of educational and social interaction effects. The solid lines describe outcomes brought about when the agents acted on the basis of their true educational levels, while dashed lines describe outcomes brought about when their educational levels had been altered.¹⁸ Lines with vertical ticks describe outcomes the agents would have brought about had they not influenced one another,

¹⁸ The straight solid line without vertical ticks thus is identical to the baseline in the previous figures.

while lines without vertical ticks describe outcomes brought about when they influence one another. Short-dashed lines report outcomes when all agents are assumed to have compulsory schooling or less, while long-dashed lines describe outcomes when all agents are assumed to have a college education.

A comparison of the various unemployment trajectories in figure 6.10 gives some insights into the relative importance of social interactions. First of all, these results suggest that a removal of the social interaction effect would have a greater influence on the unemployment level than a change in the educational levels of the unemployed. If all individuals had a college education or, expressed slightly differently, if all individuals were able to leave unemployment as fast as the college-educated could, these analyses suggest that the number of unemployed individuals during an average week would have been about 9 per cent lower than in the baseline scenario with actual education. This should be compared with what would have happened had the social interaction effect been eliminated. Such a change would have reduced the number of unemployed during an average week by approximately 11 per cent.¹⁹ A similar conclusion can be drawn from the fact that the baseline setup generates higher unemployment levels than the setup which assumes atomistic agents with compulsory education or lower only.²⁰

The results furthermore show that the combined educational and social interaction effects can be most substantial. The sharpest reduction in the unemployment level is brought about when the social interaction effects are eliminated *and* all individuals are able to leave unemployment as fast as the college-educated could (see the long-dashed line with vertical ticks). This experimental setup brings about 19 per cent fewer unemployed during an average week.

On the basis of these comparisons we therefore conclude that the social interaction effects are at least on a par with the educational effects.²¹ Performing analyses like these allows us not only to state that

¹⁹ That social interactions have a greater impact is indicated by the fact that the solid line with vertical ticks is below the long-dashed line without vertical ticks for most of the period.

²⁰ This can be seen by comparing the solid straight line with the short-dashed line with vertical ticks. During an average week, the number of unemployed was about 6 per cent lower in the latter scenario.

²¹ It should also be noted that this way of assessing the magnitude of the educational effect somewhat overestimates its true unique effect. The effects we have reported combine a direct educational effect and an indirect social interaction effect. That is, changes in the educational effects lead to changes in unemployment levels and these lead to further changes in the unemployment level because of the social interaction effect. The experimental treatment which assumes that all agents have a college education will lead to 691

a social inter
state with so
and, in this
for the socia
claims, we b
analyses in a

Con

The lack o
research tha
a large part
little or no c
often more
Coleman (C
integration
able metho
ual to the
sociologica
theories fo
on individu

In this cl
arriving at
approach a
ories and
summarize

- 1 Start wi
logic of
in orde
that is,
outcom
we hav
explana
- 2 For en
import
that th

fewer un
ence on
would b
compuls

be outcomes brought about when
ed lines report outcomes when all
ory schooling or less, while long-
all agents are assumed to have a

loyment trajectories in figure 6.10
importance of social interactions.
a removal of the social interaction
on the unemployment level than a
unemployed. If all individuals had
ghtly differently, if all individuals
ast as the college-educated could,
nber of unemployed individuals
een about 9 per cent lower than
ucation. This should be compared
the social interaction effect been
ave reduced the number of un-
approximately 11 per cent.¹⁹ A
n the fact that the baseline setup
s than the setup which assumes
cation or lower only.²⁰

t the combined educational and
ubstantial. The sharpest reduction
about when the social interaction
s are able to leave unemployment
l (see the long-dashed line with
p brings about 19 per cent fewer

we therefore conclude that the
on a par with the educational
se allows us not only to state that

is indicated by the fact that the solid line
ine without vertical ticks for most of the
aight line with the short-dashed line with
nber of unemployed was about 6 per cent

ng the magnitude of the educational effect
. The effects we have reported combine a
interaction effect. That is, changes in the
ployment levels and these lead to further
f the social interaction effect. The experi-
s have a college education will lead to 691

a social interaction process might have been of importance, but also to
state with some confidence that such processes actually were at work
and, in this case, that they most likely were of considerable importance
for the social outcome we sought to explain. Being able to make such
claims, we believe, is of utmost importance for the future of agent-based
analyses in an empirically oriented discipline like sociology.

Concluding discussion

The lack of integration between sociological theory and sociological
research that Merton so often brought to our attention still characterizes
a large part of the discipline. Theorists who specialize in theory still have
little or no contact with empirical research, and empirical researchers are
often more concerned with statistical than with sociological theory.
Coleman (1986b) argued that one important reason for this lack of
integration is that we have neither a well-specified theory nor a depend-
able methodology for making the transition from the level of the individ-
ual to the level of the social. As a result, sociological theory and
sociological research often appear mismatched. While most sociological
theories focus on social phenomena, most quantitative research focuses
on individual-level phenomena.

In this chapter we have advocated the use of so-called ECA models for
arriving at a closer integration of theory and research. The modelling
approach adopted here seeks to closely integrate mechanism-based the-
ories and empirical research, and the core of the approach can be
summarized in the following way:

- 1 Start with developing a stylized agent-based model that explicates the
logic of the mechanism assumed to be operative. Simulate the model
in order to examine *generative sufficiency* (Epstein and Axtell 1996),
that is, make sure that the model can generate the type of social
outcome to be explained. If the model exhibits generative sufficiency,
we have a mechanism-based explanation of the outcome, but the
explanation has not yet been empirically verified.
- 2 For empirical verification, use relevant data to examine the most
important bits and pieces of the causal machinery in order to verify
that the mechanism actually works as postulated.

fewer unemployed individuals during an average week if the agents interact and influ-
ence one another. Had they not interacted with one another, the corresponding figure
would have been 551. The comparable figures for the set up where the agents have
compulsory education or lower are 401 and 320.

- 3 Examine generative sufficiency when the agent-based model has been modified in the light of (2) and after controls for likely confounders have been introduced.

Only when our explanatory account has passed all of these three stages can we claim to have an empirically verified mechanism-based explanation of a social outcome.

In t
soc
tory
of
an
an
res
so
an

W
b
M
a
C
c
s
c